



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2012

Computergenerierte Figuren in Benjamin Button und Avatar

Flückiger, Barbara

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-66112>

Book Section

Published Version

Originally published at:

Flückiger, Barbara (2012). Computergenerierte Figuren in Benjamin Button und Avatar. In: Segeberg, Harro. Film im Zeitalter "Neuer Medien II" : Fernsehen und Video. München, Paderborn: Fink Verlag, 109-137.

Mediengeschichte des Films VIII · Segeberg (Hrsg.)

Mediengeschichte des Films

Herausgegeben von Harro Segeberg
in Zusammenarbeit mit Knut Hickethier, Corinna Müller
und dem Metropolis-Kino Hamburg.

BAND 8: Film im Zeitalter Neuer Medien II
Digitalität und Kino

BAND 1: Die Mobilisierung des Sehens
Zur Vor- und Frühgeschichte des Films in Literatur und Kunst

BAND 2: Die Modellierung des Kinofilms
Zur Geschichte des Kinoprogramms zwischen Kurzfilm und Langfilm

BAND 3: Die Perfektionierung des Scheins
Das Kino der Weimarer Republik im Kontext der Künste

BAND 4: Mediale Mobilmachung I –
Das Dritte Reich und der Film

BAND 5: Mediale Mobilmachung II –
Hollywood, Exil und Nachkrieg

BAND 6: Mediale Mobilmachung III –
Das Kino in der Bundesrepublik Deutschland als Kulturindustrie
(1950 – 1962)

BAND 7: Film im Zeitalter Neuer Medien I –
Fernsehen und Video

Film im Zeitalter Neuer Medien II

Digitalität und Kino

Mediengeschichte des Films
Band 8

Herausgegeben von
Harro Segeberg

Wilhelm Fink

Gedruckt mit freundlicher Unterstützung
der Hamburgischen Wissenschaftlichen Stiftung

Umschlagabbildung:

Gerhard Muche und Harro Segeberg nach Motiven aus *Metropolis* (1926)
Avatar (2009), *Welt am Draht* (1973) und *Jurassic Park II Making of* (2005).

Bibliographische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie; detaillierte bibliografische Daten
sind im Internet über <http://dnb.d-nb.de> abrufbar.

Alle Rechte, auch die des auszugsweisen Nachdrucks, der fotomechanischen
Wiedergabe und der Übersetzung, vorbehalten. Dies betrifft auch die Vervielfältigung und Übertragung einzelner Textabschnitte, Zeichnungen oder Bilder
durch alle Verfahren wie Speicherung und Übertragung auf Papier, Transparente,
Filme, Bänder, Platten und andere Medien, soweit es nicht §§ 53 und 54 UrhG
ausdrücklich gestatten.

© 2012 Wilhelm Fink Verlag, München
(Wilhelm Fink GmbH & Co. Verlags-KG, Jühenplatz 1, D-33098 Paderborn)

Internet: www.fink.de

Einbandgestaltung: Evelyn Ziegler, München
Printed in Germany
Herstellung: Ferdinand Schöningh GmbH & Co. KG, Paderborn

ISBN 978-3-7705-5327-3

Inhalt

Einleitung

Harro Segeberg: Antimimetische Mimesis. Zur Medialität und Digitalität des Kinofilms	9
--	---

Teil I Strategien der Digitalisierung

Rüdiger Maulko Mimesis und Anthropologie des Digitalen: Synthetischer Fotorealismus im Kino	29
Martin Doll Entzweite Zweifelt? Zur Indexikalität des Digitalen	57

Teil II Effekte der Digitalisierung

Klaus Kohlmann Illumination und Material – technische Komponenten des computergenerierten Films	89
Barbara Flückiger Computergenerierte Figuren in <i>Benjamin Button</i> und <i>Avatar</i> . Technik und Ästhetik	109
Jens Eder, Jan-Noël Thon Digitale Figuren in Kinofilm und Computerspiel.	139
Markus Kuhn Digitales Erzählen? Zur Funktionalisierung digitaler Effekte im Erzählkino	183

Teil III
Praktiken der Digitalisierung

Jan Distelmeyer Machtfragen. Home Entertainment und die Ästhetik der Verfügung	225
Franziska Heller <i>Prettier than Ever</i> . Die digitale Re-Konstruktion von Filmgeschichte und ihre Versprechen	253
Heinz Hiebler Digital Tools. Filmanalyse und Filminterpretation im digitalen Zeitalter.	275
Filmregister	307
Personenregister.	311
Zu den Autorinnen und Autoren.	315

Computergenerierte Figuren in *Benjamin Button* und *Avatar*

Technik und Ästhetik

Seit den frühen Tagen der Computeranimation gilt die Konstruktion von menschlichen oder menschenähnlichen Figuren als heiliger Gral der emergenten Technologie. So tauchten erste Exemplare von dreidimensionalen, voll animierten Figuren bereits zu Beginn der 1980er Jahre in Werbefilmen oder Musikvideos auf. Aber erst rund zehn Jahre später war die Entwicklung so weit fortgeschritten, dass computergenerierte Figuren in den Kinospießfilm integriert werden konnten, in *Terminator 2* (James Cameron, USA 1991), *The Lawnmower Man* (USA 1992, Brett Leonard) sowie ein digitales Double in *Jurassic Park* (USA 1993, Steven Spielberg). Nochmals rund ein Jahrzehnt später traten in *Final Fantasy* (Hironobu Sakaguchi et al., USA/Japan 2001) erstmals Figuren auf, die damals zwar als fotorealistisch eingestuft wurden, aber dennoch beim Publikum keine positiven Reaktionen hervorriefen. Zu reduziert und künstlich wirkten sie, zu hölzern und ebenmäßig. Erst mit Gollum aus *Lord of the Rings* (Peter Jackson, NZ/USA 2001-2003) gelang es, eine positiv besetzte und überzeugende Figur zu schaffen, die sowohl von Publikum wie Fachwelt akzeptiert wurde. Seither sind mehrere weitere Filme entstanden, in denen solche Figuren, teilweise sogar in Großaufnahme, zu sehen sind.

Sowohl mit der historischen Entwicklung wie auch mit technologischen, ästhetischen und narrativen Aspekten solcher computergenerierten Figuren habe ich mich in meiner Monografie *Visual Effects, Filmbilder aus dem Computer*¹ ausführlich auseinandergesetzt und dort eine Reihe von grundlegenden Einsichten in die Probleme ihrer Konstruktion und ihrer Rezeption entwickelt. Im vorliegenden Text werde ich zunächst auf diese

¹ Flückiger, Barbara: *Visual Effects. Filmbilder aus dem Computer*. Marburg 2008. Vgl. besonders das Kapitel „Körper“, S. 417-467, mit zahlreichen Abbildungen in Farbe. Diese Monografie ist eine redigierte Version meiner Habilitationsschrift an der Freien Universität Berlin (2007) und basiert auf einem Forschungsprojekt, das von 2004 bis 2006 vom Schweizerischen Nationalfonds gefördert wurde (siehe Synopsis auf http://zauberklang.ch/vfx_css.html, wo sich auch eine englische Übersetzung des gesamten Kapitels zu den digitalen Figuren zum freien Download findet).

Einsichten zurückgreifen und sie in kondensierter Form nochmals darlegen, um mich anschließend in zwei Fallstudien zu *Avatar* (James Cameron, USA 2009) und *The Curious Case of Benjamin Button* (David Fincher, USA 2008) mit dem derzeitigen State-of-the-Art zu beschäftigen. Immer geht es mir dabei um den Zusammenhang zwischen Technologie, Ästhetik und Narration.

Grundprobleme der digitalen Figurenkonstruktion

Warum ist es so schwierig, eine überzeugende digitale Figur zu konstruieren?

Einer der Hauptgründe dafür liegt in unserer lebensweltlichen Erfahrung, denn in unserem Alltag müssen wir seit frühester Kindheit lernen, die gestischen und mimischen Signale der Menschen in unserer Umwelt zu entschlüsseln, um erfolgreich kommunizieren zu können. Deshalb haben wir ein unglaublich feines Sensorium noch für unwichtig erscheinende Details entwickelt. Es ist sogar zulässig anzunehmen, dass die Wahrnehmung der menschlichen Gestalt und besonders des menschlichen Gesichts wohl jener Bereich ist, in welchem wir in der Regel über die feinste Ausdifferenzierung verfügen.

Grundproblem 1: Verknüpfung von Erscheinungsbild und Verhalten

Nun ist es so, dass in unserer Erfahrung alle diese Details in einem systemischen Zusammenhang integriert sind. Im Unterschied dazu werden computergenerierte Figuren in einem hochgradig arbeitsteiligen Prozess aus einer Vielzahl von isolierten Details zusammengesetzt, und es gestaltet sich äußerst schwierig, diese Details in einen übergreifenden Zusammenhang zu stellen und sinnvoll miteinander zu verknüpfen. Zur analytischen Diskussion des Sachverhalts habe ich vorgeschlagen, die Figuren unter den Aspekten Erscheinungsbild und Verhalten zu untersuchen, wobei sich das Erscheinungsbild weiter ausdifferenzieren lässt in die Komponenten Form und Oberfläche.² Während das Verhalten in den Kompetenzbereich der Animationsabteilung gehört, wird die Form mit einer 3D-Software modelliert und die Oberflächeneigenschaften sind durch *Shader* definiert. Shader beschreiben die Reaktionen von Oberflächen auf das einfallende Licht.

² Flückiger: *Visual Effects*, S. 433.

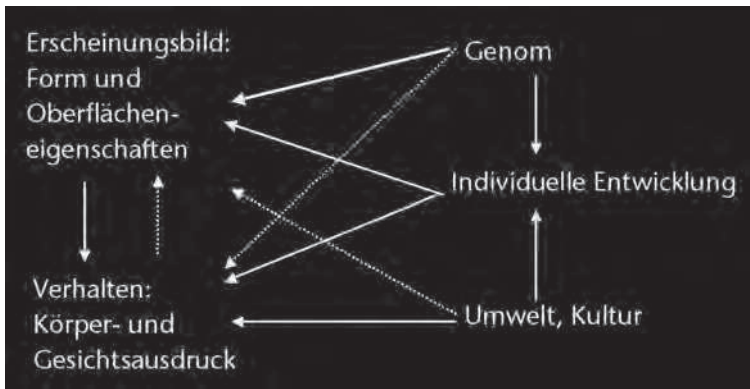


Abb. 1: Das dynamisch-interaktionistische Persönlichkeitsmodell

In der lebensweltlichen Realität bestehen zwischen Erscheinungsbild und Verhalten enge Verbindungen, die sich mit dem *dynamisch-interaktionistischen Paradigma* der Persönlichkeitspsychologie fassen lassen;³ es geht von einer engen Wechselwirkung zwischen Person und Umwelt aus und berücksichtigt somit sowohl genetische als auch umweltbedingte Faktoren der Entwicklung, die einander gegenseitig bedingen (vgl. Abb. 1). Es ist nach wie vor anspruchsvoll, eine solche natürlich wirkende Interaktion zwischen allen Eigenschaften einer künstlich geschaffenen Existenz herzustellen.

Grundproblem 2: Modellieren von Komplexität

Computergenerierte Objekte tendieren aus prinzipiellen Gründen dazu, unorganisch und unterkomplex auszufallen. Denn weil jedes Detail von Grund auf definiert werden muss, fehlen zufällige Variationen, welche die allzu starren Muster aufbrechen und unvorhersehbare Schwankungen hinzufügen. Alles in diesen Bildern muss geplant werden. Zwar gibt es prozedurale Verfahren, welche auf stochastischen oder fraktalen Algorithmen aufbauen, um pseudo-zufällige Muster zu erzeugen. Aber diese Verfahren eignen sich nicht für die Erzeugung von menschlichen oder menschenähnlichen Lebewesen. Wie ich in meinen grundsätzlichen Überlegungen zu den Modellbildungsprozessen, die im Zentrum von computergenerierten Bildern stehen, herausgearbeitet habe, ist Komplexität immer das Resultat

³ Asendorpf, Jens B.: *Psychologie der Persönlichkeit*. Berlin 2005, S. 84 ff.

eines Prozesses, also einer Geschichte.⁴ Und diese Geschichte muss in computergenerierten Figuren doppelt vorhanden sein: einerseits als eine Konzeptionsgeschichte, welche mittels einer Backstory den Designprozess steuert, der Figur also eine Herkunft zuweist, die über den unmittelbaren erzählerischen Kontext hinausgeht, und andererseits als eine Entstehungsgeschichte jedes einzelnen computergenerierten Exemplars, die im Schreiben und Umschreiben von Codes besteht. Wie ich zeigen werde, haben sich deshalb verschiedene Aufzeichnungsverfahren etabliert, welche Komplexität sozusagen aus der äußeren Welt abgreifen und in die computergenerierte Welt importieren.

Grundproblem 3: Interaktion der digitalen Figur

Anders als im reinen Animationsfilm werden die digitalen Figuren, von denen in diesem Text die Rede ist, in eine Live-Action-Umgebung eingefügt, müssen also sowohl mit den anderen, von Schauspielern verkörperten Figuren sowie mit dem Terrain und den Requisiten interagieren. Das bringt eine Reihe von Problemen mit sich und erfordert verschiedene Maßnahmen, die idealerweise schon in der Vorproduktion angelegt werden. Denn die digitale Figur muss ja irgendwie am Set vertreten sein, damit die Schauspieler darauf reagieren und ihr Timing anpassen können. Im einfachsten Fall sind es Kartontafeln, die ungefähr andeuten, wo die Schauspieler die Blicke hinwenden müssen. Vielseitiger und besser etabliert ist ein Ansatz, den ich mit dem Begriff *Proxy* fasse, also ein Stellvertreter am Set, der die Figur vertritt und damit auch eine physische Interaktion inklusive Berührungen ermöglicht. Denn Berührungen – das haben meine Untersuchungen gezeigt⁵ – sind die wohl beste Grundlage für die Glaubwürdigkeit und Präsenz der digitalen Figuren. Außerdem sind sie ein wichtiger Pfeiler für die emotionale Partizipation der Zuschauer, welche die Figur immer in Beziehung zum Ensemble wahrnimmt und deren Verhalten entsprechend bewertet. Der Proxy-Ansatz verlangt jedoch eine hoch entwickelte Tracking- und Retuschentechnik, es sei denn, dass mit einer computergesteuerten *Motion-Control*-Kamera gedreht wird, die dann sowohl die Raumdaten wie auch den *Clean Pass* ermöglicht – einen Aufnahmedurchgang ohne Schauspieler, der die Hintergründe für die Retusche liefert.⁶

⁴ Flückiger, *Visual Effects* (vgl. Anm. 1), Kapitel „Abbildung“, S. 275-356.

⁵ Flückiger, *Visual Effects*, Abschnitt „Berührungen, Blicke, physische Interaktion, Konsequenzen“, S. 249-256.

⁶ Vgl. Flückiger, *Visual Effects*, Abschnitt „Raum- und Bewegungsanpassung: Motion Control und Tracking“, S. 239-249.

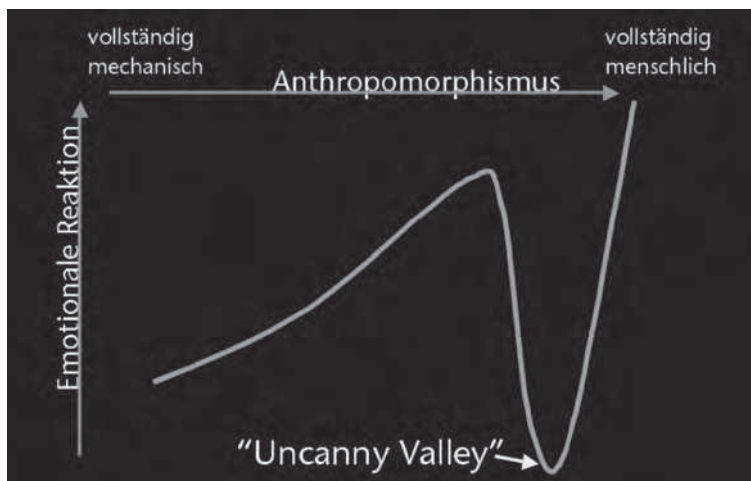


Abb. 2: Die Theorie des Uncanny Valley

Zwei Modelle zur Evaluation von digitalen Figuren

Die populäre Theorie des „Uncanny Valley“

Seit einigen Jahren diskutieren Visual-Effects-Spezialisten ein Modell, das zu Beginn der 1970er Jahre ursprünglich im Kontext der Robotik entstanden ist: die populäre Theorie des „Uncanny Valley“ (japanisch *bukimi no tani*) von Masahiro Mori.⁷ Mori hat dabei folgende Regeln in der emotionalen Bewertung von künstlichen Figuren beobachtet und beschrieben: Je menschenähnlicher eine solche Figur – zum Beispiel ein Roboter – wirkt, desto positivere Emotionen löst er aus. Wenn die Figur aber nicht ‚wirklich‘ menschlich, sondern nur *fast* menschlich erscheint, tritt ein distanzierender Effekt ein, den Mori als „Uncanny Valley“ bezeichnet hat (Abb. 2). Erst wenn die Figur als vollständig menschlich wahrgenommen wird, lässt sich dieser Effekt überwinden und die Emotionen fallen uneingeschränkt positiv aus. Zur Illustration dieses Sachverhalts zieht Mori eine prothetische Hand heran, die äußerlich gesehen zwar überzeugend aussieht, sich bei einer Berührung aber kalt, weil plastifiziert anfühlt und damit unwillkürlich einen Schauer auslöst. Ganz ähnlich verhält es sich mit digitalen Figuren. Als Paradebeispiel gilt *Final Fantasy*, das bei der Rezeption diesen befremdlichen Effekt hervorgerufen hat und entsprechend gescheitert ist.

⁷ Mori, Masahiro: „Bukimi no tani. (The Uncanny Valley)“ In: *Energy*, 1970, Nr. 7.

Die Theorie des „Uncanny Valley“ ist aber nur bedingt universell. Einerseits gibt es Figuren, die nur fast vollständig menschlich wirken und dennoch nicht Befremden hervorrufen. Andererseits – und das ist aus methodischer Sicht fast problematischer – muss die Theorie im Einzelfall empirisch überprüft werden, denn ob der Effekt eintritt oder nicht, kann strittig sein und eventuell auch intersubjektiv variieren, wie sich am Beispiel von *Avatar* zeigen lässt.

Modell der Distanz

Als Alternative habe ich deshalb das Modell der Distanz vorgeschlagen. Es geht von der These aus, dass sich die Aspekte Erscheinungsbild und Verhalten in einer ähnlichen Distanz zu einer als Standard gesetzten fotorealistischen Abbildung befinden sollten. Wenn also eine Figur sehr stilisiert ist, sollte sie entsprechend stilisiert animiert werden; eine Regel übrigens, die schon in Disneys Regelkatalog unter der Nr. 10 („exaggerate“) figurierter. Der Begriff Distanz bezieht sich entsprechend seiner Verwendung in der Prototypentheorie auf die Abweichung von einem als Norm definierten Standardwert.

Zur Darstellung dieses Sachverhalts habe ich eine Matrix entwickelt, in welche sich verschiedene digitale Figuren projizieren lassen (Abb. 3). Dabei ist die feine, aber essenzielle Linie zu beachten, die zwischen einer fotorealistischen und einer stilisierten Darstellung gezogen werden muss, denn ob und wie diese Linie überschritten wird, ist nach meiner Hypothese entscheidend für die Rezeption. So ist in *Final Fantasy* eine klare Trennung zwischen fotorealer Oberfläche und einem stilisierten, weil in seinem Ausdrucksrepertoire deutlich reduzierten Verhalten zu beobachten. Im Gegensatz dazu sind die Disney-Figuren eindeutig in der stilisierten Domäne angesiedelt. Gollum aus *Lord of the Rings* (2001-03) schließlich ist ein anderer Fall. Wie meine größere Fallstudie gezeigt hat, wirken Verhalten und Erscheinungsbild hauptsächlich fotorealistisch und organisch mit einigen Abweichungen, die als fantastisch oder zumindest ungewöhnlich einzustufen sind, so die großen Ohren und Augen oder seine Fortbewegungsart. Diese Abweichungen stören jedoch nach meiner Hypothese das Gleichgewicht der Darstellung nicht, weil sie gleichermaßen in Verhalten und Form vorkommen. Sie sind eher wie Satelliten zu verstehen.⁸

⁸ Vgl. Flückiger, *Visual Effects* (vgl. Anm. 1), Abschnitt „Von der Fallstudie Gollum zu grundsätzlichen Überlegungen“, S. 451-461.

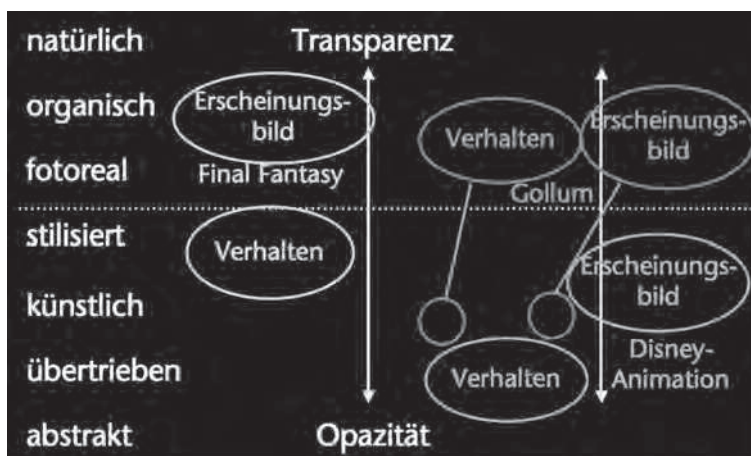


Abb. 3: Das Modell der Distanz

Fallstudie 1: *The Curious Case of Benjamin Button* (2008)

Es gilt nun also in der Fallstudie zunächst zu *Benjamin Button* zu analysieren, wie die Figur konstruiert worden ist, ob und wenn ja wie sie die zuvor benannten Grundprobleme meistert und schließlich wie sie in den Modellen zur Evaluation abschneidet.

Mit *Benjamin Button* ist unter der Regie von David Fincher ein Film entstanden, der sich nicht leicht in etablierte Kategorien einordnen lässt. Es ist ein Film, der in der Tradition des pikaresken Romans steht. Entsprechend schildert er die Lebensreise eines armen, aber bauernschlaugen Helden von der Geburt bis zum Tod und kann damit in Verwandtschaft zu *Forrest Gump* (Robert Zemecki, USA, 1994) gesetzt werden. Die erzählerische Anordnung unterscheidet sich davon aber durch ein fantastisches Element, denn der Protagonist wird als alter Mensch geboren und verjüngt sich zunehmend. Diese Konstellation hat besonders schmerzhafteste Konsequenzen für sein Liebesleben, denn es gibt nur eine kurze Zeitspanne, in welcher er und seine Liebste ein ähnliches Alter teilen und somit zueinander finden können.

Es galt also, den Protagonisten in verschiedenen Lebensaltern darzustellen, eine Herausforderung, die frühere Filme klassisch und überzeugend mit Prothetik und Maske gelöst haben, so wie für Dustin Hoffmann in *Little Big Man* (Arthur Penn, USA 1970), F. Murray Abraham in *Ama-*



Abb. 4: Head Replacement in *Benjamin Button*

deus (Milos Forman, USA 1984) oder Marlon Brando in *The Godfather* (Francis Ford Coppola, USA 1972). Nur weist dieser Ansatz ein grundlegendes Problem auf, das Visual Effects Supervisor Eric Barba folgendermaßen schildert: „The problem with old-age makeup is that it is additive, [...] whereas the aging process is reductive. You have thinner skin, less musculature, everything is receding.“⁹ Ein weiterer traditioneller Lösungsansatz sind Darsteller unterschiedlichen Alters, die eine Rolle in verschiedenen Lebensabschnitten verkörpern. Diesen Ansatz wollte Fincher vermeiden, denn er stellt einen Bruch in der Figurenidentität dar. Zusätzlich wurde die Anforderung verschärft durch die Dissonanz zwischen wachsendem Körper und altem Aussehen, und deshalb wollte Fincher eine Lösung mit einer computergenerierten Figur wagen. Allerdings ist die jeweils altersgerechte Version von Brad Pitt nur teilweise computergeneriert, denn der dazu gehörende Körper stammt von einem menschlichen Darsteller, während der Kopf vollständig digital erzeugt und mittels *Head Replacement* auf den Darstellerkörper gesetzt wird (Abb. 4). Um die Anordnung zu testen, realisierten Fincher und das Team von Digital Domain einen Werbefilm für „Orville Redenbacher’s“ Popcorn,¹⁰ der auf YouTube zu Recht als „creepy“ bezeichnet und im Internet als „Orville Deadenbacher“ zirkulierte, eben weil die Figur ganz klar ein Opfer des Uncanny Valley war.¹¹ „On the Orville Redenbacher spot, we learned

⁹ Duncan, Jody: „The Unusual Birth of Benjamin Button.“ In: *Cinefex*, 2009, Nr. 116, S. 70-99 und 118, hier S. 72.

¹⁰ <http://www.youtube.com/watch?v=Fc4p213Zg8>

¹¹ Zur Problematik der digitalen Rekonstruktion von verstorbenen Personen, die hier ebenfalls eine Rolle spielt, habe ich an der Tagung NECS in Lund am 26.6.2009 den Vortrag „Digital Reconstruction of Deceased Persons“ gehalten. Es ist dies ein

everything *not* do on *Benjamin Button*,⁶ Fincher elaborated. „We learned that we needed to minimize the amount of keyframe animation [...] what I wanted was a process by which you could literally xerox an actor's performance.“¹² Finchers Statement passt exakt in meine Überlegungen zum Mangel an Komplexität, denn Keyframe-Animation – also die Animation von Schlüsselposen, wie sie seit jeher in der Zeichenanimation üblich war – tendiert eben als Modellbildungsprozess dazu, sehr vereinfachte Resultate zu liefern, denen die Eigenheiten und Irregularitäten einer individuellen Existenz fehlen.

Animation des Verhaltens

Für die Animation – und das bedeutet eben: für die Darstellung des Verhaltens – wurden deshalb Aufzeichnungsprozesse gewählt, welche Brad Pitts mimische Muster auf das computergenerierte Gesicht projizieren sollten. Dazu kamen zwei Techniken zum Einsatz, die ineinander greifen bzw. aufeinander aufbauen.¹³ Zunächst wurde mit Mova's Contour, einem volumetrischen Aufzeichnungsverfahren,¹⁴ eine Datenbank von dreidimensionalen mimischen Informationen von Brad Pitt aufgezeichnet. Dazu hat man sein Gesicht mittels eines Schwamms mit einem phosphoreszierenden Make-up eingefärbt und mit mehreren Kameras aufgenommen (Abb. 5). Das Make-up dient dazu, ein kontrastreiches Muster auf die Haut zu applizieren. Denn Mova's Contour basiert auf der computergetützten Analyse von bis zu 10'000 Punkten im Raum, das heißt der Position von individuellen Pixeln im dreidimensionalen Koordinatensystem.¹⁵ Im Unterschied zu Performance-Capture-Verfahren, welche mit Markern oder farbigen Punkten im Gesicht arbeiten, ist dieser sogenannte Pixel-Flow wesentlich genauer, denn es wird jede kleinste Bewegung erfasst, da-

Thema, das über eine ausgesprochen ethisch-moralische Komponente verfügt und die Angst vor einem posthumanen Kino schürt.

¹² Duncan, „The Unusual Birth of Benjamin Button“ (vgl. Anm. 9), S. 75. [Herv. i. O.]

¹³ Die folgenden Darstellungen basieren auf mehreren Quellen. Neben dem erwähnten *Cinefex*-Text von Jody Duncan sind es drei Präsentationen, die vom 5. bis 8. Mai 2009 an der Tagung fmx 09 (<http://09.fmx.de/start.php?lang=E&navi=1&page=pages>) in Stuttgart vorgestellt wurden, nämlich zur Animation von Steve Preeg, Animation Supervisor, und Jonathan Litt, CG Supervisor, Digital Domain; Patrick Davenport, Executive Producer, Image Metrics; Paul Debevec, Professor für Computer Grafik an der University of Southern California USC.

¹⁴ Volumetrisch bezieht sich auf die Verteilung der Punkte im Raum.

¹⁵ Eine anschauliche Visualisierung des Verfahrens findet sich auf <http://www.mova.com/flash/> [Zugriff am 18. März 2011].



Abb. 5: Movia's Contour als volumetrisches Verfahren

mit auch die Mikro-Expressionen, die wesentlich für das Verständnis der mimischen Kommunikation sind.

Mit Movia's Contour entstand eine Datenbank mit 170 sogenannten Blend-Shapes, also Grundposen. Das waren in diesem Fall mimische Grundeinheiten, die gemäß den Regeln des Facial Action Coding Systems FACS von Paul Ekman et al. systematisiert wurden.¹⁶ Zusammenfassend wurde also in diesem ersten Prozess das Gesicht in verschiedenen Posen ohne zeitliche Komponente erfasst. Als problematisch an dieser Technik erwiesen sich die fehlenden Informationen für die Augen und das Innere des Mundes, also Zähne und Zunge, die später von mehreren Modellierern in einem Monate dauernden Prozess von Hand hinzugefügt werden mussten.¹⁷

Für den nächsten Schritt, eben die Aufzeichnung der Bewegung in der Zeit, kam ein bild-basiertes Verfahren zum Einsatz, nämlich Image Metrics' Performance Capture, das ebenfalls den Pixel-Flow analysiert. Vier hochauflösende Viper-Kameras nahmen Brad Pitts mimische Performance aus unterschiedlichen Perspektiven auf (Abb. 6). Zunächst studierte Brad Pitt die am Set aufgenommenen Einstellungen der Proxies,¹⁸ die

¹⁶ FACS: System zur Beschreibung des emotionalen Ausdrucks auf dem menschlichen Gesicht; beruht auf der Analyse des Gesichtsausdrucks anhand der darunter liegenden Muskelstruktur. Ekman, Paul; Friesen, W.: *Facial Action Coding System. A Technique for the Measurement of Facial Movement*. Palo Alto 1978. Vgl. zur Anwendung von FACS in der Gesichtsanimation Flückiger, *Visual Effects*, S.446.

¹⁷ Preeg, fmx 09 (vgl. Anm. 13).

¹⁸ Es kamen drei Schauspieler zum Einsatz, die Benjamin Button in drei verschiedenen Lebensaltern spielten (Duncan, „The Unusual Birth of Benjamin Button“, S.79)



Abb. 6: Aufzeichnung der Mimik mit Image Metrics

mit einer blauen Haube versehen Benjamin Button spielten. Die Animatoren fürchteten, dass Brad Pitt die Aufnahmen gefühlte 150'000 mal studieren müsste, bis er einigermaßen in der Lage wäre, das Timing in allen Nuancen zu erfassen, aber tatsächlich schaffte er es nach rund zweimaligem Schauen.¹⁹ Man muss sich diesen Prozess so ähnlich vorstellen wie eine optische Nachsynchronisation (*visual automated dialogue replacement, ADR*), ein Konzept übrigens, das auch bei *Avatar* Anwendung fand, wie ich später ausführen werde. Wie immer bei mehr oder weniger automatischen Aufzeichnungssystemen, sind die Resultate weit davon entfernt, sofort perfekt auszufallen. „Every shot felt a little dead and required a lot of massaging by hand.“²⁰ So stellten die Animatoren fest, dass kleinste Verschiebungen der Position des Augenlids im Bereich von 1 mm zu dramatischen Veränderungen des Ausdrucks führten, eben weil jeder Zuschauer ein Experte im Dekodieren des Gesichtsausdrucks ist.

Schließlich hat man das digitale 3D-Modell zusätzlich mit einem physikalisch-basierten Animationssystem, Dynamik flexibler Körper (*soft body dynamics*) genannt, ausgestattet, um die Verformungen der gealterten Haut, die besonders markant schwabbeln sollte, und die Faltenbildung regelbasiert zu erstellen.

Erscheinungsbild: Form und Haut

Für die Form hat man zunächst von Brad Pitts Kopf einen Gipsabdruck hergestellt, der als Grundlage für drei Silikonmodelle diente. Diese Silikonmodelle wurden von Hand angefertigt und entsprachen drei verschiedenen Lebensaltern – 60, 70 und 80 Jahre. Sie wurden bemalt und mit Hautdetails wie Poren und Haaren versehen. Es wurde also in einer manu-

¹⁹ Preeg, fmx 09 (vgl. Anm. 13).

²⁰ Preeg, fmx 09.

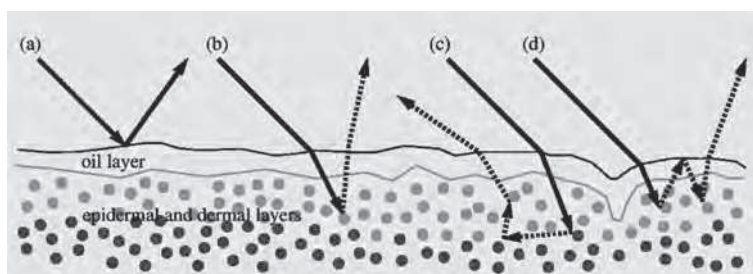


Abb. 7: Subsurface Scattering

ellen Art und Weise eine Extrapolation von Brad Pitts möglichem Alterungsprozess vorgenommen. Nicht direkt davon, sondern von einer Gipsvorstufe dieser Modelle hat Digital Domain einen *3D Scan* erstellt, um das Modell in die digitale Domäne zu importieren. Dabei wird das Modell von einem Laserstrahl rundherum abgetastet und so Punkt für Punkt die Oberfläche erfasst, der es aber noch an allen Details fehlt, die schließlich im Materialisierungsprozess als *Shader* appliziert wird. Diese Shader sind als Netzwerke organisiert, in welchen mehrere Elemente die Materialeigenschaften definieren, also die Farbverteilung, kleinräumige geometrische Variationen oder das Reflexionsverhalten sowie Aspekte der Lichtbrechung, der Diffusion, der Lichtbeugung und volumetrische Effekte, um nur einige wenige zu nennen.

Bis heute ist menschliche Haut eines der anspruchvollsten Materialien für die Darstellung in computergenerierten Bildern, und zwar besonders wegen ihrer Transluzenz, also ihrer halbtransparenten oder durchscheinenden Materialität. Denn die Lichtstrahlen dringen durch die Oberfläche in die tieferen Hautschichten ein, wo sie in komplexer Weise gestreut werden und dabei die Farbwerte des Gewebes und der Blutgefäße annehmen (Abb. 7). Verzichtet man auf dieses sogenannte *Subsurface Scattering*, wirkt die Haut wie Gips.

Ein Shader für Haut setzt sich – vereinfacht dargestellt – aus folgenden Komponenten zusammen: Texture Maps, welche in Form von Bildern die Farbverteilung definieren; Displacement Maps für die kleinräumigen Oberflächenvariationen der Haut wie Poren und Falten; Albedo Maps, welche die diffus reflektierenden Anteile der Haut erfassen, die zur Verweißlichung führen; Specularity Maps für die Glanzreflexionen; Reflectance Maps für die Farbreflexionen sowie Texture Maps der unter der äußeren Haut liegenden Farbinformationen für das Subsurface Scattering (Abb. 8).

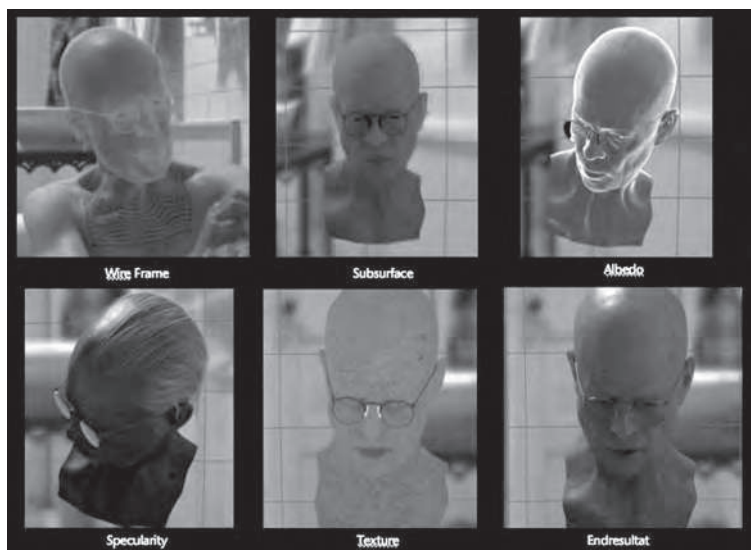


Abb. 8: Shader-Elemente

Dieser Shader wurde mit Light-Stage-Aufnahmen²¹ von den Silikonmodellen abgeglichen, die ein nicht vollständig realistisch wirkendes Resultat lieferten, da das Ausgangsmaterial ja aus Silikon bestand, aber doch eine sehr gute Vorlage darstellten.

Fazit

Es dürfte klar sein, wie hybrid und verschränkt sich der Konstruktions- und Animationsprozess gestaltete. Zwar bilden Aufzeichnungsverfahren ein wesentliches Gerüst in diesem System – und dazu möchte ich in der analogen Domäne den Gipsabdruck zählen, immerhin vergleichen sowohl Roland Barthes als auch André Bazin die Fotografie mit einem Abdruck –,

²¹ Die *Light Stage* – an der University of Southern California unter der Leitung von Paul Debevec entwickelt – ist eine Lichtbühne, in welcher der Schauspieler in der Mitte platziert, von allen Seiten mit steuerbaren Lichtdioden beleuchtet und von mehreren Kameras aufgenommen wird. Es ist also ein fotografisches Aufzeichnungsverfahren, welches die Lichtreflexion direkt erfasst. Damit lassen sich alle normalerweise in einem Shader separat definierten und verknüpften Informationen integriert aufnehmen (siehe <http://gl.ict.usc.edu>)

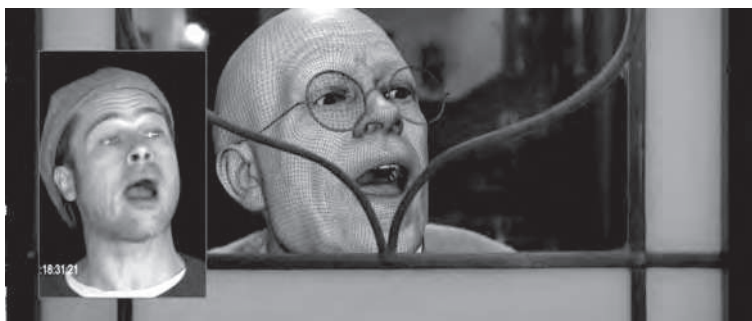


Abb. 9: Brad Pitt und Benjamin Button

aber an vielen Stellen sind es Prozesse der Modellbildung sowohl bei einzelnen Shadern, bei der Animation mit der Dynamik flexibler Körper als auch im Renderprozess, der vollständig auf Modellbildung beruht. Die organisch wirkende Komplexität der Figur ist ohne Zweifel dem dominanten Einsatz von Aufzeichnungsverfahren zuzuschreiben.

Referenz der Aufzeichnung war immer Brad Pitt. Damit ist die Figurenkonsistenz schon zu weiten Teilen gesichert. In der gealterten Version von Benjamin Button ist eindeutig Brad Pitt zu erkennen (Abb. 9), und das ist nicht zuletzt als ein Resultat der Technik einzustufen, aber eben nicht nur. Wie die Darstellungen insbesondere der Animation, aber auch der Materialisierung und der Form gezeigt haben, bilden selbst fortgeschrittene technische Verfahren nicht mehr als das Rohmaterial, das erst durch die manuelle Bearbeitung sowie die geschickte Verknüpfung der unterschiedlichsten Bausteine einen einheitlichen und überzeugenden Eindruck hervorruft.

Das Problem der Interaktion mit den anderen Figuren, den Objekten und dem Raum wurde mit dem Einsatz von Proxies, deren Körper im Film zu sehen sind, elegant gelöst. Nicht nur ist damit die physische Interaktion gegeben, sondern die Körperanimation als Ganzes entfällt und damit ein weiteres Problem der Erzeugung von digitalen Figuren. Typischerweise findet sich im Film die obligate Berührung des computergenerierten Gesichts, welche die digitale Domäne in die analoge Welt integriert. Neben der körperlichen Interaktion liegt ein weiterer Faktor für die erfolgreiche Integration der digitalen Figur in der ästhetischen Kohärenz, die einerseits auf eine sehr gute Materialisierung der Haut mit allen Details, andererseits aber auf eine ausgefeilte Übertragung der Lichtsituation am Set mittels bild-basierter Beleuchtung zurückzu-



Abb. 10: Benjamin Button in Großaufnahme

führen ist.²² Damit wird eine sehr natürlich wirkende Beleuchtungssituation auch in der digitalen Domäne möglich. Mit einem adaptiven Rendering in 20 Durchgängen, ließ sich die Figur im Compositing extrem flexibel in die Live-Action-Umgebung einpassen.²³ Das Ergebnis ist jedenfalls sehr beeindruckend ausgefallen (vgl. Abb. 10)

Schließlich ist Benjamin Button noch in die anfangs erwähnten Modelle einzuordnen. Man kann davon ausgehen, dass die Figur das Uncanny Valley überwunden wurde. Es sind nur sehr wenige kurze verunsichernde Momente zu beobachten, die Zweifel aufkommen lassen, und diese Momente werden immer durch die Krisenzonen Augen und Mund hervorgerufen, die nach wie vor am schwierigsten zu bewältigen sind. Im Modell der Distanz schließlich lässt sich Benjamin Button eindeutig dem fotorealistischen Bereich zuordnen. Anders als Gollum und auch anders als eine andere gelungene digitale Figur, Davy Jones aus *Pirates of the Caribbean: Dead Man's Chest* (Gore Verbinski, USA 2006), der mit seinen Tentakeln im Gesicht deutlich verfremdet erscheint und sich damit dem Vergleich mit einem Menschen weitgehend entzieht, sieht Benjamin Button sehr normal aus. Dennoch ist er eben ein Sonderling. Einige merkwürdige Entgleisungen des Gesichts erscheinen narrativ motiviert. Jeden-

²² Bildbasierte Beleuchtung (*image-based lighting*): ein Verfahren, bei welchem die Beleuchtung aus High-Dynamic-Range-Fotografien einer realen Umgebung abgetastet und in die computergenerierte Szene importiert wird (Vgl. Flückiger, *Visual Effects*, [vgl. Anm. 1, S. 164 ff.]). Eines der Hauptprobleme dieses Verfahrens war bisher der beschränkte Bewegungsspielraum (Vgl. Flückiger, *Visual Effects* [Anm. 1], S. 166). Wie Jonathan Litt an der fmx 09 ausführte, hat das Team von Digital Domain dieses Problem mit Nuke, einer Compositing-Software, gelöst.

²³ Litt, fmx 09 (vgl. Anm. 13).

falls erlaubt die Geschichte einen größeren Spielraum. Und es handelt sich um eine alte Figur, deren Haut nicht mehr so elastisch und auch nicht mehr so transparent ist. Auch das hat die Arbeit erleichtert, aber das soll die überaus großartige Leistung des Teams von Digital Domain nicht schmälern.

Fallstudie 2: *Avatar*

James Camerons Science-Fiction-Film sprengte im Winter 2009/2010 alle Box-Office-Rekorde. Das ist Indiz genug für das Funktionieren der digitalen Figuren, denn wie sonst sollten sich Millionen von Zuschauern rund um den Globus emotional so stark berühren lassen? Tatsächlich griffen im Kino gestandene Männer in den entscheidenden Momenten zu den Taschentüchern. Es entstand ein regelrechter Hype um den Film, der als revolutionärer Meilenstein der Visual Effects gefeiert wurde. Mit der fantastischen Welt Pandora jedenfalls war ein computergeneriertes Universum in nie zuvor gesehener Komplexität und schierer Quantität entstanden, mit fluoreszierenden Lebewesen und Pflanzen, mit riesigen, mythisch aufgeladenen Bäumen, mit schwebenden Inseln und fantastischen Tieren, und nicht zuletzt bevölkert von Massen von digital konstruierten, fremdartigen Eingeborenen namens Na'vi und den Avataren, gentechnisch hergestellten Hybriden zwischen Mensch und Na'vi, die sich von Menschen steuern ließen.

Wie in der Fallstudie zu *Benjamin Button* wird es hier nun darum gehen, die Entstehung dieser digitalen Figuren in *Avatar* zu analysieren²⁴ und basierend auf den eingangs dargelegten Problematiken kritisch zu diskutieren.

Seinen eigenen Ausführungen zufolge hatte James Cameron die Idee zu *Avatar* schon Mitte der 1990er Jahre.²⁵ Wie in *Benjamin Button* wären tra-

²⁴ Als Quellen dienen die umfangreiche Darstellung in der Fachzeitschrift *Cinefex*: Duncan, Jody: „The Seduction of Reality.“ In: *Cinefex*, 2010, Nr. 120, S. 68-146; Holben, Jay: „Mauro Fiore, ASC Helps James Cameron Envision Avatar, a 3-D Science-fiction Adventure That Combines High-definition Video and Motion Capture.“ In: *American Cinematographer*, Bd. 91 (2010), Nr. 1, Januar S. 32-47, die Präsentationen am 7. Mai 2010 an der Tagung fmx 10 von Shawn Dunn, Animation Technical Supervisor, und Stephen Rosenbaum, Visual Effects Supervisor, Weta Digital (siehe <http://2010.fmx.de/effects.E.224.html>, Zugriff am 26.5.2011).

²⁵ Duncan, „The Seduction of Reality“ (vgl. Anm. 24), S. 70. Spiegel, Simon: „Wie James Camerons Avatar den Zuschauer in eine fremde Welt entführt“. In: Mamczak, Sascha/Jeschke, Wolfgang (Hrsg.): *Das Science Fiction Jahr 2010*. München 2010, S. 361-388. Ein Scriptment mit einer frühen Version von James Cameron

ditionelle Ansätze denkbar gewesen, die Geschichte umzusetzen, nämlich eine Kombination von Make-up und Skalierung, um die größeren Figuren in Relation zu den Menschen zu setzen. Cameron jedenfalls überlegte sich, digitale Figuren mit Aufnahmen eines Regenwalds zu kombinieren. Zu jener Zeit aber war die Technik zu wenig ausgereift, um eine digitale Figur zu schaffen, die als Träger von Emotionen einen Spielfilm hätte ausfüllen können. Gemäß der Fama ließ Cameron das Projekt deshalb erstmal ruhen, wandte sich anderen Projekten wie *Titanic* zu und wartete darauf, dass sich die Technik weiterentwickelte.

Animation des Verhaltens

Schon früh bestand der Plan, mit einer erweiterten Form von Motion Capture, nämlich Performance Capture zu arbeiten. Performance Capture ist ein integrierter Ansatz zur simultanen Aufzeichnung von grobmotorischen und mimischen Bewegungen, der erstmals für Robert Zemeckis' *The Polar Express* (USA 2004) Verwendung fand.²⁶ Dieses Verfahren geht von der richtigen Annahme aus, dass sich die Mimik in Abhängigkeit von den Körperbewegungen verändert und dass deshalb die früheren Techniken, die Mimik und Gestik getrennt behandelt haben, keine überzeugenden, systemisch verknüpften Bewegungsmuster von Gesicht und Körper liefern können. „Our goal in using performance capture,“ noted *Avatar* producer and longtime Cameron collaborator Jon Landau, „was [...] to preserve the actor – because what a great actor does and what a great animator does are antithetical to each other. A great actor withholds information.“²⁷ Es ging also im Kern wie bei *Benjamin Button* genau darum, einen natürlich wirkenden Stil zu entwickeln anstelle einer akzentuierten Stilisierung in der Tradition des Animationsfilms.

Und wie in *Benjamin Button* wurde ein bildbasiertes Verfahren mit Pixel-Flow gewählt: Die Idee bestand darin, die Schauspieler mit einer Helmkamera in Standardauflösung zu bestücken und diese Aufnahmen des Gesichts zur Basisinformation für die Animation zu erklären. Auf der Motion-Capture-Bühne waren zudem High-Definition-Kameras installiert, welche die Szene als Ganzes erfassten und damit weiteres Referenzmaterial für die Animation lieferten. Während der Motion-Capture-Ses-

findet sich auf <http://www.docstoc.com/docs/14294813/Avatar-Scriptment-by-James-Cameron>.

²⁶ Vgl. Flückiger, *Visual Effects* (vgl. Anm. 1) S. 449 ff. sowie den Abschnitt „Motion Capture“, S. 145-153.

²⁷ Duncan, „The Seduction of Reality“ (vgl. Anm. 24), S. 70.



Abb. 11: James Cameron mit der virtuellen Kamera

sions in den Giant Studios, die mehrere Monate dauerten,²⁸ setzte James Cameron eine virtuelle Kamera ein (Abb. 11). Als virtuelle Kamera bezeichnet man üblicherweise den Blickpunkt, aus dem eine computer-generierte Szene dargestellt wird, weil dieser Blickpunkt dem einer Kamera entsprechen würde. Im Fall von *Avatar* hat James Cameron einen beweglichen Monitor, dessen Bewegung im Raum unmittelbar erfasst wurde, wie eine Kamera verwendet. Damit legte er die einzelnen Einstellungen der Szenen bereits fest und konnte in Echtzeit die Performance der digitalen Figuren sowie deren Integration in die computergenerierte Umwelt überprüfen, allerdings in einer niedrigen Auflösung und mit einer bescheidenen Ästhetik der gerenderten Bilder, die an frühe Computerspiele erinnert (Abb. 12).

Außerdem ging es bei dieser Technik darum, sofort einen Eindruck von den geplanten Einstellungen zu bekommen, was ohne Zweifel die Zusammenarbeit sehr erleichterte und auch für die Schauspieler von Nutzen war, denn auf der Motion-Capture-Bühne arbeiten sie in einem abstrakten, äußerst reduzierten sinnlichen Vakuum, in dem alles nur angedeutet ist, weil es gilt, Verdeckungen durch Objekte möglichst zu vermeiden, damit die Bewegungsdaten lückenlos aufgezeichnet werden können. Ein weiterer Vorteil war es für Cameron, dass die Schauspieler anders spielten, wenn sie eine Beziehung zur Kamera aufbauen und nicht nur allgemeine Daten für ein indifferentes Kamera-Array liefern mussten:

²⁸ In Holben, „Mauro Fiore“ (vgl. Anm. 24), ist die Rede von 18 Monaten, in Duncan, „The Seduction of Reality“, S. 124 (ebd.), werden 6 Monate genannt.

Abb. 12: Echtzeitvisualisierung
in *Avatar*



„Actors act differently depending on the camera,‘ noted [digital effects supervisor Nolan] Murtha, ‚if you are doing a wide master of a scene, they’re going to be a little more grand in their motion than they would be on a close-up.“²⁹

Ähnliche Echtzeitvisualisierungen, sogenannte *On-set Previz*, sind aber schon in früheren Produktionen anzutreffen, so das Encodacam-System aus dem Broadcastbereich unter anderem in *A.I. Artificial Intelligence* (Steven Spielberg, USA 2001) und in *I, Robot* (Alex Proyas, USA 2004) sowie einiges früher ein videobasiertes System in *Who Framed Roger Rabbit* (Robert Zemeckis, USA 1988).³⁰ Zemeckis hatte diese Idee weiterverfolgt und insbesondere für *Beowulf* (USA 2007) eingesetzt.

Zu den Ergebnissen der Bilder aus der Helmkamera ist anzumerken, dass diese Bilder – anders als das Image-Metrics-System aus *Benjamin Button* – keine verlässlichen Bewegungsdaten liefert. Das hat zwei Gründe. Erstens braucht es eine sehr hochauflösende Bildqualität, damit die Pixel-Flow-Analyse funktioniert; denn ohne hohe Auflösung hat man ähnliche Filterprobleme wie mit Markern.³¹ Zweitens hat man nur eine Perspektive von vorne statt mehrerer Aufnahmewinkel zur Verfügung, und dies mit den Verzerrungen eines Fischaugenobjektivs. Deshalb hat man – wie in *Benjamin Button* – die Aufnahmen nachsynchronisiert mit

²⁹ Duncan, „The Seduction of Reality“ (vgl. Anm. 24), S. 120.

³⁰ Flückiger, *Visual Effects* (vgl. Anm. 1), S. 237 f.

³¹ Weil nicht jeder Punkt exakt erfasst wird, müssen die Bewegungen in den Zwischenbereichen berechnet werden.

einem System, das sich analog zum Automated Dialogue Replacement (ADR) Facial Performance Replacement (FPR) nannte.³² Dabei wiederholten die Schauspieler ihre Mimik sitzend vor einem Monitor und wurden von mehreren Kameras aufgenommen. Mit diesem Prozedere hat man allerdings einen wesentlichen Vorteil von Performance Capture verloren, nämlich die Integration von Körperbewegung und mimischem Ausdruck: „So we were uncoupling the facial performance from the physical performance.“³³ Geblieben ist von der ursprünglichen Idee also in erster Linie das Referenzmaterial, das die Helmkamera geliefert hat.

Wie in *Benjamin Button* hat man für *Avatar* das Facial Action Coding System FACS von Ekman und Friesen verwendet, jedoch mit einem anderen Ansatz. Im Unterschied zur Blend-Shape-Technik nämlich von *Benjamin Button*, welche nur die Oberflächenveränderungen erfasst, wurden die Figuren mit einem robusten und ausgereiften Gesichtsmuskelsystem ausgestattet, das die Mimik genau nach den 44 mimischen Einheiten steuerte, die Ekman et al. definiert und mit den sechs Basisemotionen Glück, Überraschung, Trauer, Angst/Furcht, Ekel/Abscheu und Zorn verknüpft hatten. Mit dieser anspruchsvollen Technik wollte man die geringe Datenmenge des Performance-Capture-Materials ausgleichen.³⁴

Da die Gesichter der Na'vi mit ihren breiten Nasen und ihren Raubtieraugen sich deutlich von menschlichen Gesichtern unterscheiden – ein Aspekt, der uns im Kontext der Uncanny-Valley-Diskussion nochmals beschäftigen dürfte –, bestand eine Schwierigkeit darin, die Daten von den Performance-Capture-Aufnahmen entsprechend zu übertragen. Ähnliches gilt für die Körperproportionen. Die Na'vi sollten ungefähr 2.50 m groß sein, was zu erheblichen Problemen bei der Skalierung der Bewegungsdaten führen kann, wie meine Untersuchungen gezeigt haben.³⁵ Denn die Bewegungsmuster verändern sich in Abhängigkeit von Alter, Größe und vor allem Masse. Tatsächlich wirken die Na'vi oftmals zu leicht, und es ist enorm schwierig, ihre Größe korrekt wahrzunehmen, wenn keine menschlichen Figuren als Vergleichsobjekte zu sehen sind.

³² Duncan, „The Seduction of Reality“ (vgl. Anm. 24), S. 119.

³³ James Cameron in Duncan, „The Seduction of Reality“, S. 119.

³⁴ Duncan, „The Seduction of Reality“, S. 137.

³⁵ Das Skalierungsproblem bei Motion Capture ließ sich schon andernorts beobachten, so bei *The Polar Express* (Robert Zemeckis, USA 2004) und *Hulk*, siehe Flückiger, *Visual Effects* (vgl. Anm. 1), S. 151 f.

Erscheinungsbild: Form und Haut

Im Entstehungsprozess des Erscheinungsbilds unterscheidet sich *Avatar* fundamental von *Benjamin Button*, denn die Na'vi und die Avatare sind frei erfundene Fabelwesen mit variierenden menschlichen Anteilen. In solchen Fällen gibt es üblicherweise einen längeren Konzeptionsprozess, der mit Zeichnungen beginnt. Die ersten Entwürfe hatte Cameron selbst angefertigt, schon da mit blauer Haut, an welcher er durch den ganzen Prozess festhielt, um die Figuren von allen etablierten grünen Monstern und Marsmännchen klar abzusetzen. Ebenfalls von Camerons Design stammten die großen Katzenaugen, die breiten Löwennasen, die großen beweglichen Ohren und die großen, schlanken, muskulösen Körper.³⁶ Es ist aber nicht überraschend, dass der Konzeptionsprozess insgesamt immer stärker zu menschenähnlichen Figuren konvergierte, denn die Anschlussfähigkeit, um einen Begriff von Niklas Luhmann zu verwenden,³⁷ ist die Basis für den Verstehensprozess und damit auch für die empathische Partizipation. Diese Erkenntnis gewannen auch die Filmemacher: „eventually we went back to a much more traditional foundation. We came to realize that if we wanted the audience to relate emotionally to these characters, there needed to be familiar touchstones.“³⁸ Um den Performance-Capture-Abgleich zu erleichtern, hat man die Mundpartie möglichst von den Schauspielern übernommen. Bei den Avataren – besonders bei Sigourney Weaver – musste die Ähnlichkeit mit der menschlichen Figur gegeben sein, zumindest im Gesicht. Dafür hat man sicherlich 3D-Scans der Schauspieler angefertigt.

An den Konzeptionsprozess schließt die Übertragung in den dreidimensionalen Raum an. Es werden meist Tonmodelle angefertigt, die man anschließend in 3D scannt, um sie in den Computer zu laden. Bei *Avatar* war es etwas anders. Zwar gab es diesen Weg, aber parallel dazu hat man die Zeichnungen in eine Software namens ZBrush von Pixologic geladen, um dort in einem Sculpting genannten Prozess 3D-Modelle im Computer zu erstellen (Abb. 13). ZBrush kam vor wenigen Jahren auf und ist heute ein sehr etabliertes Verfahren, um komplexere organisch wirkende Objekte oder Figuren zu erstellen.³⁹ Es ist eine Art dreidimensionales

³⁶ Duncan, „The Seduction of Reality“ (vgl. Anm. 24), S. 75.

³⁷ Luhmann, Niklas: *Die Realität der Massenmedien*. Wiesbaden 2004 (zuerst 1995), S. 35 ff. Vgl. dazu meine Überlegungen zur filmischen Fiktion: Flückiger, *Visual Effects* (vgl. Anm. 1), S. 282–289.

³⁸ Duncan, „The Seduction of Reality“ (vgl. Anm. 24), S. 75.

³⁹ Kingslien, Ryan (2004): *ZBrush 2.0 Review* (= <http://vfxworld.com/?atype=articles&cid=2310>).



Abb. 13: ZBrush-Modell von Neytiri, rechts im Bild

Malprogramm, mit dem sich im Raum Formen, Texturen, kleinräumige Variationen und weitere Materialeigenschaften erzeugen lassen.⁴⁰

Lebensgroße Modelle der Figuren hat das Stan Winston Studio hergestellt, das wohl etablierteste Studio in diesem Sektor, das schon für die Dinosaurier in *Jurassic Park* zuständig war. Solche Modelle haben den Zweck, das Design für die Kommunikation zu veranschaulichen und als Stand-ins am Set Referenzmaterial sowohl für die Interaktion mit den Schauspielern als auch für die Beleuchtung zu bieten. Je ein weibliches und männliches Modell der Na'vi in Lebensgröße hat man in 3D gescannt und daraus die großen Massen von Na'vi-Figuren abgeleitet.⁴¹

Das Stan Winston Studio war aber auch in den Prozess der Materialisierung der Hauteigenschaften involviert, angefangen mit der Bestimmung des Blautons: „The battle was coming up with a blue that would still look believable as flesh. There are blue-colored fish and birds, but no mammals with blue that we could reference.“⁴² Zusätzlich mussten die Streifenmuster mit leuchtenden Farben entworfen und festgelegt werden, welche auf die Körper und Gesichter aufgetragen werden sollten. Aber wie Shawn Dunn⁴³ ausführte, waren diese Streifen schwierig zu handhaben, denn sie konfigurierten mit der Lesbarkeit des Gesichtsausdrucks. Im

⁴⁰ Siehe Website von Pixologic (= <http://www.pixologic.com/home.php>) sowie verschiedene Videos auf YouTube, die den Prozess veranschaulichen.

⁴¹ Duncan, „The Seduction of Reality“ (vgl. Anm. 24), S. 109.

⁴² Effects Supervisor John Rosengrant in Duncan, „The Seduction of Reality“, S. 109.

⁴³ Dunn, fmx 10 (vgl. Anm. 24).

Vergleich mit den Na'vi hatten die Avatare einen etwas helleren Hutton und ein ausgeprägteres menschliches Aussehen.⁴⁴ Sie sollten auch ihren menschlichen Pendants gleichen, um den Wiedererkennungswert zu stützen. So behielt der Dr. Grace Augustine-Avatar die typische schmale Nase der Schauspielerin Sigourney Weaver. Außerdem war die Korrespondenz zwischen steuerndem Schauspieler in der Performance Capture und digitaler Figur notwendig zur Übertragung der Daten. Besonders die Mund- und Augenpartie waren kritisch, weshalb hier die größte Übereinstimmung anzustreben war. Wie sich an einem Vergleich zwischen der Na'vi-Häuptlingstochter Neytiri und der für sie agierenden Schauspielerin Zoë Saldana leicht sehen lässt, stimmt die Übereinstimmung in erster Linie für den Mund, während die Augen in Form, Farbe und Position im Gesicht deutlich davon abweichen, so wie der Rest des Gesichtes. Neben der auffällig veränderten Nase und den hochstehenden, beweglichen Ohren hatten die Na'vi nur angedeutete Augenbrauen, was die Gesichtsanimation deutlich erschwerte, denn die Veränderungen der Augenbrauen sind wesentlich für den Ausdruck.⁴⁵

Vielleicht würde man annehmen, dass blaue Haut leichter zu materialisieren ist als die übliche menschliche Haut, weil die Referenz fehlt und daher ein großer Anteil an kritischer Begutachtung durch den Zuschauer entfallen könnte. Offenbar war das Gegenteil der Fall, wie die Praktiker berichten. Das Problem bestand darin, dass blaue Haut nie natürlich aussieht und besonders bei orangefarbenem Licht wie Feuerschein einen gräulichen, widerwärtigen Farbton produzierte, der mit entsprechenden Maßnahmen beseitigt werden musste, indem man die spektrale Reflexion der Na'vi-Haut mit einem Algorithmus verschob. Zudem ebnete der Blauton die kleinen üblichen Unebenheiten der Haut aus. Daher mussten Mitarbeiter von Weta-Digital herhalten: „Courageous Weta crew members subjected themselves to extremely close-up photography [...]; but texture artists found that they had to exaggerate [...] the freckles and other imperfections: ‚Our texture maps looked a lot more ratty than the real person's face.‘“⁴⁶ Auch diese Beobachtung stimmt mit meiner Einsicht überein, dass es notwendig ist, das allzu glatte Erscheinungsbild von CGI (Computer Generated Imagery) mit Störungen anzureichern und aufzurauen, damit es organisch und natürlich wirkt.⁴⁷ Des Weiteren ist

⁴⁴ Duncan, „The Seduction of Reality“ (vgl. Anm. 24), S. 75.

⁴⁵ Dunn, fmx 10 (vgl. Anm. 24).

⁴⁶ Visual Effects Supervisor Guy Williams in Duncan, „The Seduction of Reality“, S. 133.

⁴⁷ Flückiger, *Visual Effects*, (vgl. Anm. 1), S. 341.

anzunehmen, dass die Differenzierungsfähigkeit des menschlichen Auges im Blauspektrum geringer ist, eben weil diese Farbe in der natürlichen Hautwahrnehmung fehlt und wir entsprechend auf keine Erfahrungswerte zurückgreifen können. Und natürlich musste das zuvor erwähnte Subsurface Scattering berücksichtigt werden, wie schon bei Hulk mit seiner grünen Haut, bei welchem sich ähnliche Probleme ergaben.⁴⁸ Dafür wählte man ein Absorptions- statt des üblichen Streuungsmodells, um die Farbtöne der Unterhaut und des Gewebes aus dem Pink-Orange-Spektrum kontrolliert durchscheinen zu lassen.

Fazit

In der Marketingstrategie von *Avatar* wird Performance Capture sehr in den Vordergrund gestellt. In einem kurzen Clip, der zum Electronic Press Kit gehörte, das den Journalisten zur Verfügung gestellt wurde, kommentiert Cameron: „Alles, was wir auf der Leinwand sehen, wird von den Schauspielern zum Leben erweckt.“ Und Sam Worthington (Jake Sully) doppelt nach: „Das Aufregendste ist, das bin wirklich Ich.“ Cameron: „Wir kriegen 100% von dem, was der Schauspieler tut.“ Sam Worthington: „Es ist meine Persönlichkeit. Das bin Ich.“ Zu sehen sind Gegenüberstellungen von Performance Capture und ausgearbeiteter Animation (Abb. 14).⁴⁹ Cameron ging sogar so weit, eine Screen-Actor's-Guild- oder Oscar-Nomination für seine Schauspieler einzufordern,⁵⁰ ein Ansinnen, das Kristin Thompson kritisch reflektiert.⁵¹

Die Stoßrichtung ist klar und sie deckt sich mit meinen Beobachtungen zu Gollum. Wie ich dort schon belegt habe, war es sinnvoll, die Figur Gollum im öffentlichen Bewusstsein der Leistung eines Schauspielers zuzuschreiben: „Aus rezeptionspsychologischer Sicht [...] war das ein cleverer Schachzug, denn indem diese Leistung einer Person attribuiert wird, erlangt die digitale Figur so etwas wie eine physische Präsenz, die weitaus konkreter ist als die abstrakte und unverständliche Tätigkeit eines Heers

⁴⁸ Flückiger, *Visual Effects*, S. 442

⁴⁹ Diese Gegenüberstellung mit Split-Screen scheint sich etabliert zu haben, siehe auch Abb. 9 für Benjamin Button.

⁵⁰ Abramowitz, Rachel: 'Avatar's' animated acting. Hollywood debates whether film work used by actors such as Zoe Saldana for computer-generated characters deserves equal recognition. Los Angeles Times, 18. Februar, 2010: <http://articles.latimes.com/2010/feb/18/entertainment/la-et-avatar-actors18-2010feb18>

⁵¹ Thompson, Kristin (2010): *Motion-capturing an Oscar*, auf David Bordwells Blog: <http://www.davidbordwell.net/blog/?p=7126>



Abb. 14: Performance Capture und animierte Version

von Animatoren.⁵² Weiter wollte Cameron natürlich auch seinen Status als Regisseur sichern, das wird aus seinem Statement im oben zitierten Artikel aus der *Los Angeles Times* deutlich, wo er betont, dass er in der Zusammenarbeit mit den Schauspielern den Ausdruck kontrolliert und seine Intention nicht durch die Animatoren verändert haben möchte, ein Befund, der von Shawn Dunn gestützt wurde: „He was very religious to the performance. Whatever was shot on the mocap stage had to be in the movie.“⁵³ Insofern geht es über reine Marketingstrategie hinaus, und hier können wir den Bogen schließen zu den Grundproblemen, namentlich zum Problem der Komplexität, die – wie erwähnt – durch die Aufzeichnungsverfahren wie Motion Capture wesentlich gestützt wird.

Mehr noch als bei *Benjamin Button* jedoch ist bei *Avatar* eine hybride Mischung am Werk, die Aufzeichnungs- und Modellbildungsverfahren miteinander kombiniert, und zwar in allen Bereichen, vom Modellieren über die Materialisierung bis hin zur Animation. So sind die Figuren mit ZBrush modelliert, aber teilweise auch von Modellen oder Schauspielern in 3D gescannt; so setzen sich die Shader aus gemalten und fotografierten sowie algorithmisch bestimmten Elementen zusammen; und schließlich wird aus der reinen Betrachtung der Bewegungsabläufe klar, dass nicht alles auf der Motion-Capture-Bühne aufgezeichnet, sondern mindestens punktuell keyframe-animiert sein musste. Im *Cinefex*-Artikel schätzt Ca-

⁵² Flückiger, *Visual Effects* (vgl. Anm. 1), S. 460.

⁵³ Dunn, *fmx* 10 (vgl. Anm. 24).

meron selbst den Anteil auf 10%.⁵⁴ Diese Schätzung dürfte zu tief liegen, denn sie bezieht sich hauptsächlich auf jene Körperteile, welche die Schauspieler nicht oder in anderer Form besitzen wie die Ohren und den Schwanz, nicht aber auf das Mundinnere, das Dunn erwähnte,⁵⁵ das schon bei *Benjamin Button* überaus anspruchsvoll war und das generell eine Krisenzone darstellt. Mit der Keyframe-Animation verbunden ist aber das schon oben erwähnte Problem der Konsistenz. Wie Animation Supervisor Richard Baneham ausführte, war das Animationsgerüst (animation rig) des Gesichts mit der ausgeklügelten Muskelstruktur entscheidend für die konsistente Mimik, denn die Animation ging durch sehr viele Hände. „But the fine nuance of the actor's performance had to be done by keyframe. The facial rig just freed the animators to concentrate on those all-important details.“⁵⁶

Immer, wenn der Aktionsradius des Menschen überschritten werden soll, wie vor allem bei den Sprüngen und den Tauchern in die Tiefe in *Avatar*, muss man auf Keyframe-Animation zurückgreifen. Damit wird aber auch ein typisches Problem computergenerierter Figuren deutlich, nämlich dass sie nicht durch die physikalischen Gesetze limitiert sind. Diese annähernd unlimitierten Möglichkeiten können dazu führen, dass die Zuschauerpartizipation einbricht.⁵⁷ Merkwürdigerweise tendieren die Körperformen der Na'vi und der Avatare dazu, unterkomplex auszufallen. Es ist jedenfalls die deutliche Tendenz festzustellen, die Körper stereotypen Geschlechtermustern unterzuordnen, der schon für die frühe Computeranimation typisch ist und als Resultat der grundsätzlichen Tendenz von computergenerierten Bildern, generische Prototypen hervorzubringen, gewertet werden muss, eben weil alles geplant und von Grund auf kreiert werden muss. Die Na'vi entsprechen außerdem einem vereinfachten und traditionell festgelegten Bild des edlen Wilden, wie sie beispielsweise in Leni Riefenstahls Fotografien der Nuba zu sehen sind. Diese unterschwellige Tendenz wird als latent rassistische und oder imperialistische Ideologie des Films teilweise sehr kritisch unter die Lupe genommen, obwohl sie der weit herum rezipierten antikolonialistischen Botschaft der Erzählung zuwiderzulaufen scheint.⁵⁸ Das Hauptziel für die Gestaltung

⁵⁴ Duncan, „The Seduction of Reality“ (vgl. Anm. 24), S. 138.

⁵⁵ Dunn, fmx 10 (vgl. Anm. 24)

⁵⁶ Duncan, „The Seduction of Reality“ (vgl. Anm. 24), S. 138

⁵⁷ Vgl. meine Diskussion des Superheldenproblems in Flückiger, *Visual Effects* (vgl. Anm. 1), S. 462 ff.

⁵⁸ <http://io9.com/5422666/when-will-white-people-stop-making-movies-like-avatar>

von Neytiri jedenfalls lautete gemäß Dunn: „People should think: ‚Wow, she’s beautiful!‘“⁵⁹

Trotz des überaus hybriden Herstellungsprozesses und einer Vielzahl von Referenzen ist die Figurenkonsistenz erstaunlich überzeugend ausgefallen, denn der Bezug zu den Proxies sowohl im Erscheinungsbild als auch im Verhalten ist deutlich und schafft – ungeachtet der signifikanten Abweichungen – eine Brücke zwischen Erscheinungsbild und Verhalten, wenn auch die Bewegungsmuster oftmals der postulierten Körpermasse nicht gerecht werden, wegen der Skalierung, die dem Verhältnis zwischen Schauspielerkörper und Avatarkörper zuzuschreiben ist, das etwa einem Maßstab 1:1.5 entspricht.

Das Problem der Interaktion der Figuren fällt wesentlich komplexer aus als in *Benjamin Button*. Verschiedene Problemzonen spielen eine Rolle. Zunächst gibt es öfter eine direkte Konfrontation zwischen Live-Action und computergenerierten Bildern, so in der Einführungsszene der Avatare. Wie beschrieben, sind das schwierige Momente, weil die beteiligten Figuren ‚ontologisch‘ gesehen verschiedene Welten bevölkern. In der Einführungsszene und einigen weiteren Szenen hat man das Problem mit Motion-Capture am Set gelöst. Das klingt wesentlich einfacher, als es ist, aber James Cameron ist – anders als von ihm postuliert⁶⁰ – nicht der Erste, der diese Idee hatte. Schon in *The Lord of the Rings* und *Pirates of the Caribbean: Dead Man’s Chest* hat man diesen Ansatz verwendet.⁶¹ Grundsätzliche Schwierigkeiten resultieren daraus, dass die Lichtsituation am Set für die Motion-Capture-Kameras alles andere als optimal ist. Im Fall von *Avatar* kam als wesentliches Hindernis das Skalierungsproblem ins Spiel, denn Sam Worthington war ja deutlich kleiner als sein Avatar. Deshalb hat man das Set teilweise ebenfalls skaliert gebaut und die Handlungen ließen sich nicht in einem Durchgang aufzeichnen, sondern mussten in zwei Teile gesplittet werden: „Typically, the crew would capture the Na’vi or Avatar character first, and then map that performance to the digital character when Cameron shot live-action.“⁶² Grundlegend für das Funktionieren dieses Systems war die Echtzeitvisualisierung, welche das Team *Simulcam* nannte und das auf einer Kombination des Motion-Capture-Systems und der Live-Action-Kamera beruhte.

Weitere Problemfelder der Interaktion mussten auf der Motion-Capture-Bühne gelöst werden. So war auch dort die Skalierung ein Thema,

⁵⁹ Dunn, fmx 10 (vgl. Anm. 24).

⁶⁰ Duncan, „The Seduction of Reality“ (vgl. Anm. 24), S. 126

⁶¹ Vgl. Flückiger, *Visual Effects*, S. 151

⁶² Duncan, „The Seduction of Reality“ (vgl. Anm. 24), S. 127

das man gelöst hat, indem man die Menschen von Kindern spielen ließ, um die Proportionen abzubilden.⁶³ Die Topografie wurde mit einfachen Bauelementen angedeutet und für die Flüge auf den drachenähnlichen sogenannten Banshees hat man bewegliche Plattformen aufgehängt, deren Positionswechsel von Hand gesteuert wurden, um die Verbindung Tier-Reiter zu schaffen. Nicht nur die vielen Objekte, sondern vor allem die zahlreichen Personen, die gleichzeitig während der Motion-Capture-Sessionen aktiv waren, brachten große Verdeckungsprobleme mit sich, die man in der Vergangenheit zu vermeiden versuchte, weil die Marker dann aus dem „Sichtfeld“ der Kameras verschwinden und ihre Bewegungspfade deshalb Lücken aufweisen. Mit einem robusten Skelett und einer fortgeschrittenen Interpolationssoftware der Giant Studios ließen sich diese Probleme offenbar lösen.⁶⁴ Entscheidend aber für die gelungene Interaktion war gemäß Dunn die zuvor beschriebene Echtzeitvisualisierung der bewegten Figuren in der computergenerierten Umwelt über die virtuelle Kamera.⁶⁵ Denn in der Regel haben die Schauspieler in der Vergangenheit völlig sinnentleert auf der kargen Motion-Capture-Bühne agiert, ohne eine Vorstellung davon zu haben, wie sich ihre Performance in das Gesamtgefüge integrieren würde.

Am Ende dieses Textes geht es nun noch um eine Diskussion von *Avatar* im Kontext der beiden Modelle. Während Cameron in verschiedenen Interviews explizit und im *Cinefex*-Artikel implizit postuliert, dass seine Figuren das Uncanny Valley überwunden haben, wäre hier sicherlich eine empirische Überprüfung notwendig. Die Differenzen in der Physiognomie sind doch sehr ausgeprägt, es gibt viele Abweichungen von einem komplett menschlichen Gesicht, die den Schluss nahe legen, dass die Figuren eben nicht völlig anthropomorph sind.

Im Modell der Distanz verhalten sich die Figuren ähnlich wie Gollum aus *The Lord of the Rings*: Sie sind hauptsächlich fotorealistisch mit einigen Abweichungen, die das Gleichgewicht nicht stören, weil sie Erscheinungsbild und Verhalten ausgeglichen betreffen. Somit scheint dies ein durchaus robuster Ansatz zu sein, überzeugende digitale Figuren zu schaffen. Denn welche Kritik man auch immer an *Avatar* formulieren mag: Diese digitalen Figuren waren in der Lage, die Massen zu begeistern und ins Kino zu locken.

⁶³ Dunn, fmx 10 (vgl. Anm. 24).

⁶⁴ Duncan, „The Seduction of Reality“ (vgl. Anm. 24), S. 99

⁶⁵ Dunn, fmx 10 (vgl. Anm. 24).

Schlussbemerkung

Ohne Zweifel zeigen die beiden Fallstudien, dass die Konstruktion und Animation von digitalen Figuren inzwischen so weit ausgereift ist, dass die Resultate zu überzeugen vermögen. Für alle drei genannten Grundprobleme – Verknüpfung von Erscheinungsbild und Verhalten, Komplexität sowie Interaktion – haben sich inzwischen Lösungsansätze etabliert. Dieser Befund entspricht meiner Feststellung, dass Komplexität immer das Resultat einer Geschichte sei,⁶⁶ und diese Geschichte ist eben auch die Geschichte der Computergrafik und -animation, die bis heute eine junge Disziplin geblieben ist und ihren eigenen kulturellen Speicher noch entwickeln muss.

Jedenfalls ist alle überprononcierte Fortschrittsrhetorik, wie sie besonders von James Cameron gerne und überaus medienwirksam zelebriert wird, fehl am Platz. Trotz der zunehmend komplexeren Ausgestaltungen der Figuren, die in *Benjamin Button* und *Avatar* zu konstatieren sind, greifen alle dort verwendeten Verfahren auf Vorläufer zurück, die mehrere Jahre früher schon Verwendung gefunden haben und die mir alle in meiner Forschung schon begegnet sind. Nicht zuletzt deshalb erweisen sich meine dort gewonnenen Einsichten sowie meine Analysen und Modelle als erstaunlich robust.

Wenn man die einzelnen Techniken im Detail verfolgt, so sieht man ein komplexes Netzwerk von Einflüssen und Erkenntnissen, die aus der akademischen Grundlagenforschung aber auch aus den Research- und Developmentabteilungen der Firmen selbst stammen. Dieses Netzwerk differenziert sich stetig, aber nicht in Sprüngen aus. Besonders im direkten Austausch mit den Akteuren – den Visual-Effects-Spezialisten und den Forschern an den Universitäten – wird deutlich, wie mühsam und anspruchsvoll sich diese Arbeit gestaltet. Immer noch befindet sich die Disziplin in einem Stadium der Bricolage, wo inzwischen viele hoch entwickelte Bausteine schon zur Verfügung stehen, wo aber nach wie vor um jede einzelne Lösung gerungen werden muss. Deshalb wird die Konstruktion und Animation von digitalen Figuren weiterhin einer der aufwändigsten, aber auch spannendsten Gegenstandsbereiche computergenerierter Bilder bleiben.

⁶⁶ Flückiger, *Visual Effects* (vgl. Anm. 1), S. 317.